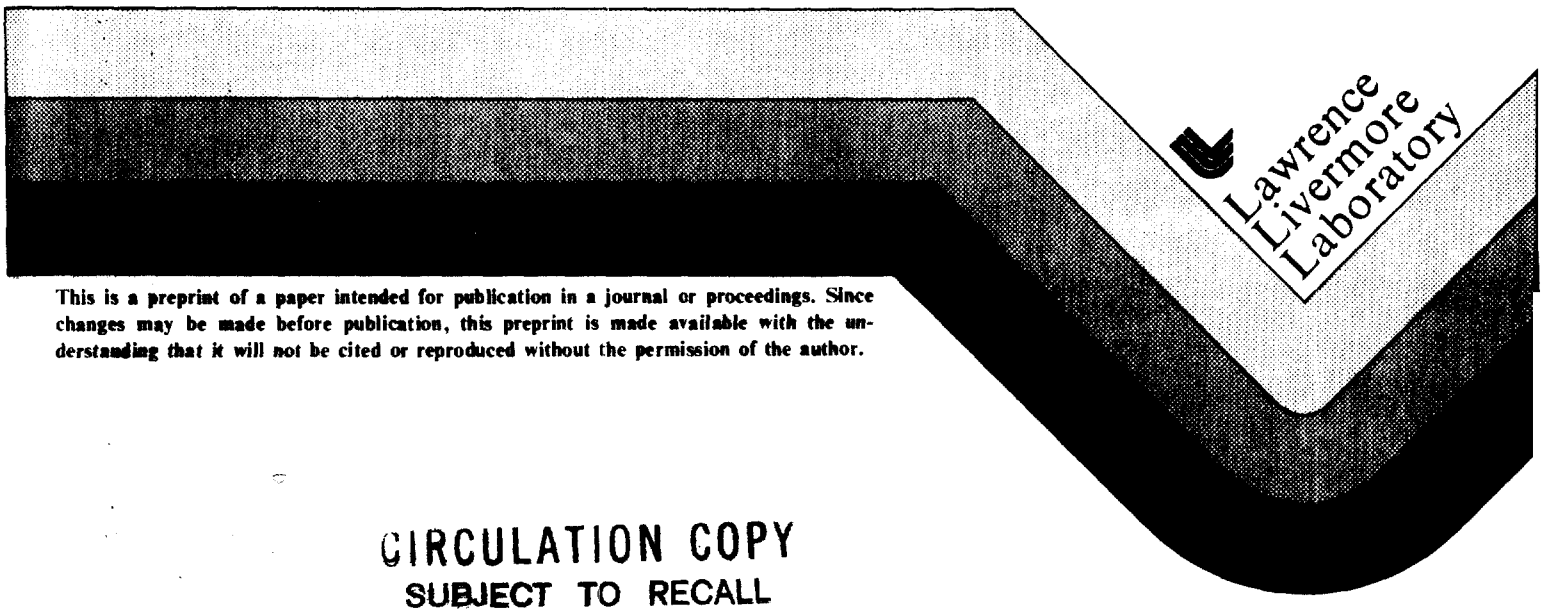
UCRL- 84281
PREPRINT

A MAXIMUM ENTROPY TECHNIQUE FOR ELECTROMAGNETIC OR SEISMIC INVERSION

R. M. Bevensee
University of California
Lawrence Livermore Laboratory
Livermore, California 94550

URSI International Symposium 1980, Munich, Germany, August 26-29, 1980

April 15, 1980



A MAXIMUM ENTROPY TECHNIQUE FOR ELECTROMAGNETIC OR SEISMIC INVERSION

R. M. Bevensee
University of California
Lawrence Livermore Laboratory
P.O. Box 808, L-156
Livermore, California 94550

ABSTRACT

We describe the application of a maximum entropy method (MEM) not to spectral analysis but rather to inversion of an underdetermined constitutive parameter (electrical conductivity, refractive index or reciprocal wave velocity) profile in a two dimensional region probed by ray data. In accord with the First Principle of Data Reduction, we maximize the probability of the distribution subject to measurement constraints and consistent totality of that parameter. An MEM algorithm is outlined and examples given of the inversion of synthetic ray data in a cell model of the earth. Computational properties of the algorithm relative to convergence, ray redundancy and noise in the ray data are described. We extend logically the MEM to include curved rays or to solve the more difficult problem of underground conductivity inversion from surface potential measurements.

We believe the examples of this particular MEM suggest an unsuspected capability for resolving two-dimensional parameter profile anomalies from a minimum of ray data. The success of our algorithm invites further analysis and application to practical field inversion problems.

*Work performed under the auspices of the U.S. Department of Energy by the Lawrence Livermore Laboratory under contract number W-7405-ENG-48.

1. HISTORY

The MEM has been applied extensively to spectral analysis since 1967 when Burg[1] presented his classic paper. The correspondence between the MEM spectral analysis and the autoregressive representation by least squares fitting of a random process was established by van den Bos.[2] Application has been made to spatial array data processing[3] and to time series analysis of signals[4] to infer the parameters of seismic events. MEM has been applied successfully to two dimensional digital image reconstruction,[5] important in radio astronomy and tomography. Our work is most closely related to the latter work of inferring the properties of an object from its noisy image, given the transmission function.

2. PROBLEM

Our model of the ground is the two dimensional cellular one shown in Fig. 1. Each of K square cells has its own (unknown) value of scalar constitutive parameter σ_k which represents conductivity, refractive index, or reciprocal wave velocity. Straight rays such as the $i^{\underline{th}}$ one are assumed to travel from a transmitter to a receiver. D_{ik} is the known path length of ray i in cell k; the measured ray datum T_i represents net attenuation, phase shift, or travel time, respectively, along the ray path. Our problem is to estimate the unknown distribution σ_k by $\hat{\sigma}_k$, $1 \le k \le K$, given the true T_i of $I < K$ rays related to the $\hat{\sigma}_k$ as

$$\sum_{k=1}^{K} D_{ik}\hat{\sigma}_k = T_i, \qquad 1 \le i \le I < K, \tag{1}$$

with or without noise in the T_i, subject to the First Principle of Data Reduction[5]: "The result of any transformation imposed on the experimental data shall in-

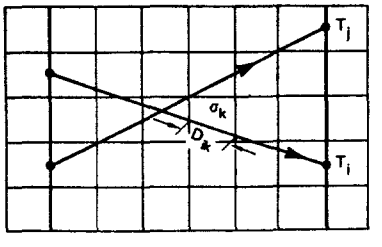


Fig. 1. Two dimensional model of the ground.

corporate and be consistent with all relevant data and be maximally noncommittal with regard to unavailable data."

3. MAXIMUM ENTROPY SOLUTION

We invoke the artifice of imagining each cell constructed of a vast number of building blocks of minute parameter $\Delta\sigma$, the $j^{\underline{th}}$ cell containing $n_j = \sigma_j/\Delta\sigma$ blocks. The totality of blocks $N = \sum n_j$ is initially regarded as fixed, though later it will be adjusted in the algorithm in accord with Eq. (1). ($\Delta\sigma$ will not appear explicitly in the solution.) We seek the most probable distribution of the n_k subject to the constraints of constant N and the measurements (1). Thus we maximize the logarithm of the total number of combinations of distinct block arrangements (entropy) subject to constraints. We use Lagrange's method of undetermined multipliers γ and β_i' to maximize instead the function F of the $\hat{\sigma}_i$,

$$F(\hat{\sigma}) = \ell n \frac{N!}{\prod_{k=1}^{K} n_k!} + \gamma \sum_{k=1}^{K} n_k + \sum_{i=1}^{I} \beta_i' \left[\sum_{k=1}^{K} D_{ik} n_k \Delta\sigma - T_i \right]. \qquad (2)$$

Differentiation relative to the n_k furnish K relations, and the other I+1 conditions required for evaluating the γ and β_i' are the constraint relations.

The straightforward solution for the $\hat{\sigma}_k = \hat{n}_k \Delta\sigma$ is as follows. The $\beta_j = \beta_i' \Delta\sigma$ must be found to satisfy the I relations, $1 \le j \le I$,

$$\sum_{k=1}^{K} D_{jk} \exp\left[\sum_{i=1}^{I} \beta_i D_{ik}\right] \Big/ \sum_{k=1}^{K} \exp\left[\sum_{i=1}^{I} \beta_i D_{ik}\right] = T_j/\hat{\sigma}_T, \qquad (3)$$

where $\hat{\sigma}_T = \sum_k \hat{\sigma}_k$ is the total estimated amount of parameter in the volume and each $\hat{\sigma}_j$ is given by

$$\hat{\sigma}_j = \hat{\sigma}_T \exp\left[\sum_{i=1}^{I} \beta_i D_{ij}\right] \Big/ \sum_k \exp\left[\sum_{i=1}^{I} \beta_i D_{ik}\right] \qquad (4).$$

One can show that any set of variations $\delta n_j = \delta\sigma_j/\Delta\sigma$ about this distribution which maintains the constraints does not change entropy S, and that $\partial^2 S/\partial n_k^2 < 0$ for all k (maximum entropy).

4. THE ALGORITHM

We investigated the properties of the MEM solution, Eq. (3)-(4), by specifying "true" σ-distributions of interest and computing for each of the synthetic ray data T_i. We made the following assumptions in solving for each MEM $\hat{\sigma}$-distribution:

(A) The true, "measured," data T_i were initially uncorrupted by noise or uncertainty.

(B) The total-parameter value $\hat{\sigma}_T$ was known to be within 10% or so of the true one σ_T.

(C) The ray path data D_{ij} were all known exactly for the straight rays.

(D) The solution should start from a pure guess of the $\bar{\beta}$-distribution (the bar denotes an array).

The algorithm computed consecutively

(i) The estimates $\hat{\sigma}_k$ from the $\bar{\beta}$ by Eq. (4).

(ii) The estimates $\hat{T}_j$ from the $\bar{\beta}$ by Eq. (3).

(iii) A correction to $\hat{\sigma}_T$ by a factor obtained from the implied constraint

$$\sum_i \beta_i \hat{T}_i = \sum_i \beta_i T_i, \quad \left(\sum_i \hat{T}_i = \sum_i T_i \text{ if } \bar{\beta} = 0\right), \qquad (5)$$

in which the left side is proportional to $\hat{\sigma}_i$ according to Eq. (3).

(iv) All the scaled $\hat{T}_i$ and $\hat{\sigma}_j$, scaled by the same factor obtained in (iii).

(v) Values of the derivatives $\partial\hat{T}_i/\partial\beta_j = \partial\hat{T}_j/\partial\beta_i$ obtained analytically from Eq. (3).

(vi) The changes $\Delta\beta_j$ necessary to improve the computed ray data $\hat{T}_i$ by solving the matrix of equations

$$T_i - \hat{T}_i = \sum_{j=1}^{I} \frac{\partial\hat{T}_i}{\partial\beta_j} \Delta\beta_j, \qquad 1 \le i \le I \qquad (6a)$$

by LU decomposition and back substitution, whereupon

$$\beta_i \leftarrow \beta_i + \Delta\beta_i, \text{ all I rays.} \qquad (6b)$$

The algorithm then returned to step (i) and repeated the loop iteratively until consecutive $\hat{\sigma}$- and $\hat{T}$-arrays after step (iii) showed convergence to three significant digits or better. $\hat{\sigma}_T$ also converged to σ_T with this accuracy in the examples.

For typical problems like those described next the algorithm generally converged on the CDC 7600 computer with the first or second choice of the initial $\bar{\beta}$-distribution (all zeroes, 0.5, or 1). Convergence when it occurred was attained in fewer than ten iterations of the loop, with a fraction of a second computing time.

Despite the vastly underdetermined nature of some of the examples, the number of rays initially specified was reducible, as indicated by the extremely small magnitudes ($< 10^{-20}$, for example) of some of the pivot elements in the triangularization of the $\partial\hat{T}/\partial\beta$-matrix in step (vi) of the loop. When the number of rays was reduced sufficiently, no extremely small pivot elements appeared and the MEM $\hat{\sigma}$- and $\hat{T}$(=T) arrays were the same as before. We cannot provide any criteria with which to test for reducibility, but its presence in practical problems as indicated by the mathematical solution is fortuitous from a data processing standpoint.

In some problems of significantly large volumes of high σ-contrast (see example 5.3) the convergence region for the starting $\bar{\beta}$-distribution was very sensitive to the T-array and to noise imposed upon it. For such problems it would be expedient to replace Eq. (6a) by a more refined, steepest-descents method of improving $\bar{\beta}$.

5. EXAMPLES

5.1 A 36-cell region (Fig. 2) probed by 20 rays, with an irregular region of $\sigma_i = 3$ contrasting slightly with the outer region of $\sigma_j = 4$. Resolution in the absence of noise in the T_i is good, inasmuch as 8 of the 12 inner cells have their $\hat{\sigma}_i$ within 0.2 of the true value of 3; 19 of the 24 outer cells are within 0.2 of the correct value of 4.

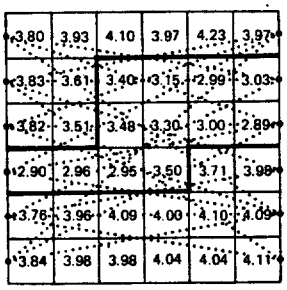


Fig. 2. MEM 20-ray reconstruction of the σ=4 outer region and σ=3 inner region shown.

5.2 A 48-cell region (Fig. 3) probed initially by 32 rays, containing an isolated sharp anomaly of $\sigma = 50$ in one cell and with $\sigma_i = 5$ in all the other cells. The rays are those shown in Fig. 4, plus the seven additional horizontal ones. The anomaly without noise is sharply revealed; only one neighbor has a $\hat{\sigma}_i$ of more than 10. Only 13 of the 47 outer cells have $\hat{\sigma}_i$ outside the range 4-6. This number of rays proved reducible; with all the horizontal rays except the top one removed, small pivot elements in step (vi) of the algorithm tended not to appear, depending on the starting $\bar{\beta}_0$ and the converged $\hat{\sigma}$- and $\hat{T}(=T)$ distributions were unchanged.

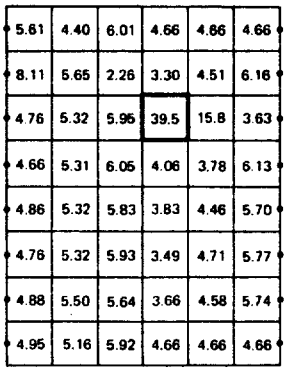


Fig. 3. MEM 25-ray reconstruction of a one-cell σ=50 anomaly, surrounded by σ=5 cells.

5.3 A 48-cell region (Fig. 4), probed by the 25 rays shown, with the 12 cells in the lower right corner of $\sigma_i = 20$ and the remainder of $\sigma_j = 5$. Without noise the corner anomaly is sharply distinguished from one block of outer cells with the MEM value of $\hat{\sigma}_i = 7.14$ and the remaining block of $\hat{\sigma}_j = 2.86$.

In practice the ray data T_i in Eq. (1) would be corrupted by error or noise and we must assess the quantitative effect upon the $\hat{\sigma}_k$. Rather than write a noise constraint explicitly into Eq. (2) we chose to create an ensemble of solutions $(\hat{T},\sigma,\bar{B})$ for a given problem, one member being the non-noise solution $(T_0,\hat{\sigma}_0,\bar{B}_0)$ and each of the others corresponding to an array of T_i all generated independently as $T_{i0}(1-f_n + 2f_n*RND)$, f_n being the fractional noise specified and RND being a random number with a uniform distribution over the interval (0,1). The relative uncertainty $\hat{\rho}_i$ in $\hat{\sigma}_i$ over this ensemble was then defined for that f_n as

$$\hat{\rho}_i = \hat{\sigma}_{i,rms}/\hat{\sigma}_{io} = \overline{(|\hat{\sigma}_i - \hat{\sigma}_{io}|^2)}^{1/2}/\hat{\sigma}_{io} . \qquad (7)$$

The following remarks refer to a 100-member noise ensemble.

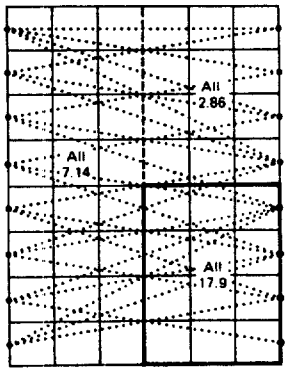


Fig. 4. MEM 25-ray reconstruction of a 12-cell anomaly of σ=20, surrounded by σ=5 cells.

The results were instructive for $f_n = 0.1$ (10% uncertainty in all the ray data, distributed uniformly). In example 5.1 with a 3:4 contrast in σ_i the $\hat{\rho}_i$-values were relatively large, being $\geq$0.2 in 31/36 of the cells. According to Fig. 2, this noise would cause most of the cells touching the σ-discontinuity boundary to have a range of uncertainty greater than the discontinuity in σ. The MEM resolution would be lost.

In example 5.2, with 25 rays defining a large discontinuity in one cell 10% noise in the ray data blurred the discontinuity negligibly. $\hat{\rho}_i$ for the $\hat{\sigma}_0 = 39.5$ cell was about .093; i.e., $\hat{\sigma} = 39.5 \pm 3.7$. The anomaly remained sharply defined. The $\hat{\rho}_i$-value for the $\hat{\sigma}_0 = 15.82$ cell next to the anomaly was 0.21, and many of the uncertainties in the other cells

lay in the range $0.4 \leq \beta_i \leq 0.6$. The cell just above the anomaly had the highest β_i of 0.79.

Example 5.3 represents an intermediate case regarding size and contrast of the anomaly. The largest β_i of the $\delta_i = 7.14$ cells was 0.80 ($\delta_i = 7.14 \pm 5.71$); of the $\delta_i = 2.86$ cells, 0.60 ($\delta_i = 2.86 \pm 1.71$); and of the $\delta_i = 17.9$ cells, 0.35 ($\delta_i = 17.9 \pm 6.27$). Thus the high-σ region remains sharply defined in the presence of 10% uniformly distributed and uncorrelated noise in all the T_i.

6. EXTENSIONS OF THE MEM

It appears that we could resolve the inverse problem with the additional constraint that the rays obey a ray-optic equation or Snell's Law at discontinuity surfaces. To avoid the difficulty of defining and computing an integrated entropy density we could leave the region discretely defined and write the constraint into the algorithm, correcting the rays according to the most recent δ-distribution between steps (iii) and (iv) in the loop. The procedure of then varying the β_i (i.e., the δ_k) at step (iv) for fixed (bent) ray paths of known D_{ij} is justified by the fact that the entropy should be maximized relative to variations $\delta\delta_i$ along fixed ray paths. This is consistent with Fermat's Principle, which says the ray travel times should be minimum relative to path variations at fixed δ.

Such a solution would conveniently neglect primary reflections and the resultant multiple rays. It would also prove convenient to ignore the information contained in the direction of approach of each ray to its receiver.

We have also applied the MEM to the difficult problem of inferring an underground layered conductivity profile from surface voltage measurements. We studied a "canonical" model 5 cells (and layers) deep and 8 cells wide in Fig. 1, with a 1-volt generator at surface node 3 and with the vertical and bottom planes at 0 volt. The data points for surface voltage measurements V_i were taken at nodes 4, 5, and 7. The problem is underdetermined by one layer because σ_1 in layer 1, say, can be specified as reference and the relative σ_i in the other four layers computed from the three surface voltages. The absolute level of the δ_i could then be inferred from a surface resistivity measurement.

The constraint relations are of the form $[\hat{V}_i(\delta) - V_i(\sigma)] = 0$ for Eq. (2) and the MEM solution is valid with $\partial\hat{V}_i(\delta)/\partial\delta_k$ replacing D_{ik} and V_j replacing T_j in Eq. (2)-(4). The best algorithm to date appears to be based on a loop which (i) computes the $\hat{V}_i$ and $\partial\hat{V}_i/\partial\delta_j$ from the most recent estimate δ of σ, (ii) updates the δ_i in the top I layers to bring $\hat{V}$ closer to V, with equations like (6), (iii) updates the $\partial\hat{V}/\partial\delta$ array again, (iv) determines the I β_i from logarithms of the equations analogous to (3), (v), obtains all the δ_k from the β_i with equations like Eq. (4), and than (vi) returns to step (i).

Preliminary numerical work indicates a capability of the MEM to distinguish a localized underground anomaly layer of higher or lower conductivity than the neighbors.

7. REFERENCES

[1] J. P. Burg, "Maximum Entropy Spectral Analysis," presented at the 37th Annual Meeting, Soc. of Exploration Geophysicists, Oklahoma City, OK, Oct. 1967.

[2] A. van den Bos, "Alternative Interpretation of Maximum Entropy Spectral Analysis," IEEE Trans. Info. Th., IT-17, 1971.

[3] R. N. McDonough, "Maximum Entropy Spatial Processing of Array Data," Geophys. 39, 1974.

[4] T. E. Landers and R. T. Lacoss, "Some Geophysical Applications of Autoregressive Spectral Estimates," IEEE Trans. Geosci. Elect., GE-15, 1977.

[5] S. J. Wernecke and L. R. D'Addario, "Maximum Entropy Image Reconstruction," IEEE Trans. Comp., C-26, 1977.